DOCUMENT RESUME

ED 069 787 TM 002 273

AUTHOR Fitzgibbon, Thomas J.
TITLE Norm Referenced and Criterion Referenced Tests from a Publisher's Point of View.
PUB DATE Sep 72
NOTE 12p.; Paper prepared for presentation at a Symposium on "The Relative Strengths of Norm Referenced and Criterion Referenced Achievement Tests (Convention of the Amer. Psychological Assn. Honolulu, Hawaii Sep. 5, 1972)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Comparative Analysis; *Criterion Referenced Tests; *Measurement Instruments; *Norm Referenced Tests; Objectives

ABSTRACT

The current popularity of criterion-referencing has been demonstrated in attacks on standardized tests and the increasing use of criterion-referenced instruments in state assessments. Much of the advocacy of the criterion-referenced test, however, seems to be less an illustration of that test's merits than an attack on the supposed deficiencies of the norm-referenced test. It is argued that norm-referenced tests are automatically biased against minorities, "do not help the learner to learn." and are not related to today's curriculum. It is pointed out that while such arguments may be true of some norm-referenced tests, criterion-referencing does not automatically correct such shortcomings. It is also pointed out that both types of tests have a place along the continuum of measurement uses and that publishers must take the lead in developing appropriate instruments for the teacher's use, whether these be called formative, criterion-referenced, or diagnostic. Finally, it is pointed out, the criterion-referenced movement can be a process which forces one to be clear about goals before one starts. (Author/CK)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED 069787

TM 002 263

# *NORM REFERENCED AND CRITERION REFERENCED TESTS FROM A PUBLISHER'S POINT OF VIEW

DR. THOMAS J. FITZGIBBON, VICE PRESIDENT
DIRECTOR, TEST DEPARTMENT
HARCOURT BRACE JOVANOVICH, INC.

MY REMARKS WILL BE MADE FROM THE POINT OF VIEW OF A MAJOR TEST PUBLISHER WHO IS DEEPLY INVOLVED IN THE SCOPE AND PROCESS OF AMERICAN EDUCATION. I CERTAINLY DO NOT PRETEND TO SPEAK FOR OTHER EDUCATIONAL TEST PUBLISHERS (I DON'T WANT TO DO SO AND I'M SURE THEY DON'T WANT ME TO), BUT I SUSPECT THAT MANY OF THE THINGS I WILL SAY ARE GENERALIZABLE TO THEIR EXPERIENCES -- IF NOT TO THEIR ATTITUDES ABOUT THESE EXPERIENCES. AND SINCE TIME IS SHORT, I HAVE CHOSEN TO MENTION ONLY A SIGNIFICANT FEW OF THE OPINIONS AND ATTITUDES GENERATED IN OUR HBJ TEST DEPARTMENT BY THE CRITERION REFERENCED MOVEMENT.

AND MOVEMENT IT IS. ITS IMPACT CAN BE FELT EVERYWHERE -- ITS MENTION IN ATTACKS ON STANDARDIZED TESTS, ITS USAGE IN STATE ASSESSMENT PROGRAMS, AND ITS CLOSE RELATIONSHIP TO INDIVIDUALIZED INSTRUCTION ARE ONLY A FEW

---

*PREPARED FOR PRESENTATION AT A SYMPOSIUM ON "THE RELATIVE STRENGTHS OF NORM REFERENCED AND CRITERION REFERENCED ACHIEVEMENT TESTS" 1972 CONVENTION OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION, HONOLULU, HAWAII, SEPTEMBER 5, 1972

## ABSTRACT

NORM REFERENCED AND CRITERION REFERENCED TESTS FROM A PUBLISHER'S POINT OF VIEW

The current popularity of criterion-referencing has been demonstrated in attacks on standardized tests and the increasing use of criterion-referenced instruments in state assessments. Much of the advocacy of the criterion-referenced test, however, seems to be less an illustration of that test's merits than an attack on the supposed deficiencies of the norm-referenced test. It is argued that norm-referenced tests are automatically biased against minorities, "do not help the learner to learn," and are not related to today's curriculum. Such arguments may indeed be true of some norm-referenced tests, but criterion-referencing does not automatically correct such shortcomings.

Too much time has been wasted insisting on the dichotomy. Both types of tests have a place along the continuum of measurement uses. Publishers must take the lead in developing appropriate instruments for the teacher's use, whether these be called formative, criterion-referenced, diagnostic, or what. We have a responsibility to furnish better interpretive materials with our instruments, to improve and extend our measurement training efforts, and to cooperate more fully with teacher-training institutions.

The criterion-referenced movement can lead to niggling specificity and rigidity, yet it can also be a process which forces one to be clear about goals before one starts and that cannot help but be a good. We shall continue our criterion-referenced endeavors with cautious optimism.



Thomas J. Fitzgibbon
Vice President and
Director, Test Department
Harcourt Brace Jovanovich, Inc.

EXAMPLES. IT IS A MOVEMENT BEHAVIORISTIC IN NATURE WHICH HAS BECOME THE DARLING OF THOSE WHO ARE DEDICATED TO EDUCATIONAL MANAGEMENT BY OBJECTIVES, E.G. THE PPBS APPROACH.

I MENTION ATTACKS ON STANDARDIZED TESTS NOT FROM A DEFENSIVE STANDPOINT -- WE CAN'T AFFORD TO BE DEFENSIVE. A TEST PUBLISHER'S ROLE SHOULD NOT BE ONE OF DEFENDER OF ONE TYPE OF TEST OR ADVOCATE OF ANOTHER, BUT RATHER ONE OF SEEKING MEASUREMENT DEVICES AND SYSTEMS WHICH WILL SERVE GENUINE EXPRESSED NEEDS. I MUST SAY THAT WHILE IN PURSUIT OF THIS GOAL WE HAVE SUFFERED MANY CHALLENGES TO THE STANDARDIZED ACHIEVEMENT TEST FROM REPRESENTATIVES OF TWO MAIN GROUPS: FIRST THE PSYCHOMETRIST OR TEST SPECIALIST WHO FEELS THAT SUCH A TEST IS TOTALLY INADEQUATE TO MEET TODAY'S PRESSING INSTRUCTIONAL NEEDS; AND SECOND, THE MINORITY GROUP REPRESENTATIVE WHO FEELS A STANDARDIZED TEST IS AUTOMATICALLY BIASED AND UNFAIR, AND THAT A CRITERION REFERENCED TEST WILL HAVE NO SUCH SHORTCOMINGS. RUNNING THROUGH BOTH OF THESE CHALLENGES IS THE FEELING THAT NORMS ARE EITHER UNNECESSARY OR DOWNRIGHT HARMFUL. IN BOTH CASES IT SOMETIMES SEEMS THAT ADVOCACY OF CRITERION REFERENCED TESTING IS LESS AN ILLUSTRATION OF ITS MERITS THAN AN ATTACK ON NORM REFERENCED TESTING.

LET'S EXAMINE THIS SITUATION IN MORE DETAIL. THE TEST SPECIALIST (OR AN EVALUATION SPECIALIST WHO MAKES USE OF TESTS) OFTEN DECLARES THAT THE STANDARDIZED ACHIEVEMENT TEST IS TOO GENERAL IN THE SWEEP OF ITS OBJECTIVES (INDEED IT IS OFTEN INFERRED THAT THERE ARE NONE AT ALL); THAT IT IS NOT RELEVANT TO TODAY'S CURRICULUM; AND THAT NORMS ARE NOT HELPFUL IN AIDING A LEARNER TO LEARN. THE FIRST DECLARATION IS IRONICALLY GENERAL IN ITS PHRASING - THE STANDARDIZED ACHIEVEMENT TEST! WHAT IS THAT? DOES THAT MEAN ALL STANDARDIZED TESTS? IT SEEMS TO ME THAT THERE ARE SOME STANDARDIZED ACHIEVEMENT TESTS WHICH HAVE BEEN VERY SPECIFIC IN THEIR OUTLAY OF OBJECTIVES TO BE TESTED FOR -- AND SOME WHICH HAVE NOT, OR AT LEAST SEEM TO HAVE NOT. IN THE LATTER CASE, I SUSPECT THERE ARE MANY OF THESE WHICH DO HAVE A RATHER COMPREHENSIVE OBJECTIVES STRUCTURE; A STRUCTURE WHICH, HOWEVER, HAS NOT BEEN MADE EXPLICIT TO THE REVIEWER, OR USER. THIS MAY OR MAY NOT BE A GOOD -- DEPENDING UPON HOW ONE LOOKS AT IT. AND HOW ONE LOOKS AT IT SO OFTEN DEPENDS NOT ONLY UPON ONE'S PEDOGOGICAL POINT OF VIEW BUT UPON FEARS THAT SPECIFICITY IN OBJECTIVE CITING (AND RESULTANT ASSIGNING OF INDIVIDUAL ITEMS THERETO) WILL LEAD TO THE BELIEF THAT THE TEST IS NOT ONLY THE FINAL CURRICULAR WORD BUT THAT IT IS ALSO DIAGNOSTIC IN NATURE. I PERHAPS HAVE DELAYED TOO LONG IN STATING THAT

I, AS WELL AS THE CRITICS, AM TALKING ABOUT SURVEY TYPE STANDARDIZED ACHIEVEMENT TESTS. IT SHOULD BE MENTIONED HERE THAT THERE ARE STANDARDIZED DIAGNOSTIC TESTS AS WELL, WHICH SEEM NOT TO HAVE BEEN DISCOVERED BY MANY CRITICS OF STANDARDIZED TESTS. FINALLY, BEFORE LEAVING THE TOPIC OF GENERALIZATION OR LACK OF SPECIFICITY OF OBJECTIVES, I MUST STATE THAT WE HAVE NOT BEEN ABLE TO FIND ANY EASY ANSWERS TO THE PROBLEM CREATED BY DEMAND FOR MORE SPECIFICITY, WHEN IT IS OFTEN ACCOMPANIED BY A REFUSAL TO DEVOTE MORE TIME TO TESTING. NOT AN UNUSUAL CIRCUMSTANCE, I MUST REPORT. INEVITABLY, NO MATTER WHAT THE TEST IS CALLED, WE ARE MOST OFTEN FORCED INTO DESCRIBING A DOMAIN AND THEN SAMPLING WITHIN THAT DOMAIN. SOME OF THE TALK ABOUT THE PENETRATING EFFICIENCY OF OBJECTIVES BASED TESTING IS DELUSORY - IT FOUNDERS UPON THE PRAGMATIC ROCK OF THE CLASSROOM SCHEDULE.

THE CHARGE THAT STANDARDIZED ACHIEVEMENT TESTS ARE NOT RELATED TO TODAY'S CURRICULUM MAY ALSO BE TRUE OR FALSE DEPENDING UPON THE TEST IN MIND -- AND JUST WHAT THE CRITIC THINKS TODAY'S CURRICULUM IS. CRITERION REFERENCED TESTING, AS A REPLACEMENT FOR NORM REFERENCED, WILL NOT GUARANTEE CURRICULAR RELEVANCY. WHAT WILL GUARANTEE IT IS TEST AUTHORS AND PUBLISHERS WHO ARE ALERT TO CURRICULUM CHANGE, WHO ARE RESPONSIVE TO IT IN TERMS OF WILLINGNESS

TO DEVOTE TIME AND MONEY TO ADJUSTMENT TO THAT CHANGE, AND FINALLY WHO ARE SKILLFUL ENOUGH AT THEIR TRADE TO DO IT IN A PROFESSIONALLY ACCEPTABLE WAY.

THE CHARGE THAT NORMS ARE NOT HELPFUL IN AIDING A LEARNER TO LEARN IS A VERY SERIOUS ONE AND SHOULD BE SUBJECTED TO VIGOROUS RESEARCH EFFORTS. FROM THE PUBLISHER'S POINT OF VIEW, THE ECONOMIC BENEFITS FROM DROPPING THE STANDARDIZATION COMPONENT OF THE NEW-TEST RESEARCH PROGRAM WOULD BE VERY, VERY SUBSTANTIAL. INDEED, OUR RELATIONSHIPS WITH SCHOOLS MIGHT BE EVEN BETTER SINCE WE ARE CONSTANTLY CALLING UPON THEM FOR COOPERATION IN EXPERIMENTATION. NEEDLESS TO SAY, THIS COOPERATION IS NOT ALWAYS EASY FOR THEM TO SCHEDULE OR TO AFFORD IN TERMS OF TIME AND TEACHER ACCEPTANCE. THE MORE ESOTERIC THE RESEARCH PROJECT, I.E., LESS RELATED TO WHAT A SCHOOL PERSON SEES AS BEING IMMEDIATELY AND DIRECTLY HELPFUL TO HIM, THE MORE DIFFICULT IT IS TO ELICIT THE NECESSARY SCHOOL-PUBLISHER COOPERATION. THE ABOVE STATEMENT, HOWEVER, IS A DIGRESSION FROM THE REAL PROBLEM. NO ONE KNOWS, TO MY KNOWLEDGE, WHETHER OR NOT NORMS DO MAKE A DIFFERENCE ON "WHAT" AND "HOW MUCH OF WHAT" A PERSON LEARNS. I AM CONVINCED, HOWEVER, THAT NORM REFERENCING, I.E., THE ERECTING OF A FRAME OF REFERENCE TO WHICH PUPIL OR SCHOOL SCORES MAY BE COMPARED, WILL CONTINUE TO BE DEMANDED. A SOCIETY SUCH AS OURS, AGRESSIVE AND STRIVING AS IT IS, WILL

INSIST UPON MAKING COMPARISONS, WITH THE INSISTENCE COMING FROM SUCH DIVERSE GROUPS AS LEGISLATORS, COMMUNITY GROUPS, AND INDIVIDUAL PARENTS. ONE MAJOR EFFORT WE SHOULD MAKE RELATES NOT TO THE ELIMINATION OF NORMS PER SE, BUT TO CONSTRUCTION OF THE NECESSARY FRAMES IN SUCH A WAY THAT THEY ARE UNDERSTANDABLE AND RELEVANT TO THE COMPARISONS WHICH ARE GOING TO BE MADE. INVIDIOUS COMPARISONS WILL ALWAYS BE WITH US; IGNORANT COMPARISONS WE CAN DO SOMETHING ABOUT. A SECOND MAJOR EFFORT SHOULD BE MADE TO BUILD ANOTHER SET OF DEVICES (CALL THEM CRITERION REFERENCED OR CONTENT REFERENCED IF YOU WILL) WHICH FOCUS MORE UPON DIFFERENCES BETWEEN WHAT A YOUNGSTER KNOWS AND WHAT SOCIETY THINKS HE SHOULD KNOW. I THINK IT IS SAFE TO SAY THAT MOST MAJOR PUBLISHERS HAVE SPENT A GREAT DEAL OF TIME AND ENERGY ON BOTH OF THESE EFFORTS IN THE LAST 3 TO 5 YEARS.

THE STATE ASSESSMENT PROGRAM INVOLVEMENT WITH CRITERION REFERENCED TESTS OFTEN INTRODUCES YET ANOTHER KIND OF THINKING ABOUT THESE INSTRUMENTS, I.E., THAT TEACHER INVOLVEMENT IN THE DEFINING OF GOALS, IN THE SELECTION OF OBJECTIVES, AND IN THE WRITING OF ITEMS TO MEASURE WHETHER THOSE OBJECTIVES HAVE BEEN ATTAINED IS AN AUTOMATIC GOOD. FOR SOME UNSETTLING REASON, MANY CRITERION REFERENCED ADVOCATES SEEM TO FEEL THAT MOST ANY ONE CAN WRITE A GOOD

ITEM IF OBJECTIVES ARE SPELLED OUT, THAT GOOD ITEMS CAN BE DEVISED IN A SHORT PERIOD OF TIME, AND THAT "PUTTING THE TEST TOGETHER" IS A RELATIVELY SIMPLE PROPOSITION. MY PREDICTION IS THAT MORE THAN ONE STATE ASSESSMENT PROGRAM WILL FOUNDER IF THESE MISCONCEPTIONS CONTINUE TO EXIST. INDEED, THE PRESENT STATE OF AFFAIRS IN MANY STATE ASSESSMENT PROGRAMS IS DISCOMFITING TO ME. CONCEIVED OFTEN IN AN ENVIRONMENT OF SUSPICION; GESTATED DURING A PERIOD WHICH SEES LEGISLATORS, SCHOOL ADMINISTRATORS, AND TEACHERS INCREASINGLY HOSTILE TO ONE ANOTHER; AND FINALLY, FORCIBLY DELIVERED BECAUSE OF UNREAL TIME CONSTRAINTS, THEY CAN CAUSE GREAT HARM TO THE CRITERION REFERENCED MOVEMENT -- IN FACT, ALL TESTING. I CANNOT LEAVE THIS STAND IN GOOD CONSCIENCE IF I DO NOT ADMONISH ALL OF US, PROFIT ORGANIZATIONS, NOT-FOR-PROFIT ORGANIZATIONS, UNIVERSITY CONSULTANTS, AND OTHERS WHO SEEK TO BECOME INVOLVED IN SUCH PROGRAMS, TO CONSIDER THEIR RESPONSIBILITY TO THOSE WHO INSTIGATE THEM AND TO THOSE WHO ARE DIRECTED TO IMPLEMENT THEM. PUT IN ANOTHER WAY, MANY OF THEIR MEASUREMENT DEMANDS ARE UNREALISTIC. THOSE OF US WHO KNOW THIS TO BE THE CASE SHOULD SAY SO AND SHOULD SUGGEST OUR ALTERNATIVES. IF THESE ALTERNATIVES ARE NOT ACCEPTED, AND IF WE CONTINUE TO FEEL THAT THE PROJECT IS IMPOSSIBLE, WE SHOULD NOT BID OUR SERVICES.

A GREAT DEAL OF TIME HAS BEEN WASTED ARGUING CRITERION vs. NORM REFERENCED TESTING. THEY ARE NOT DICHOTOMOUS CONDITIONS; RATHER, THEY EXIST ALONG A CONTINUUM OF USES OF MEASUREMENT INSTRUMENTS. TO USE THE TERM "VERSUS" PRESUPPOSES AN ADVERSARY SITUATION WHICH CANNOT EXIST IF WE ARE TO HAVE A RANGE OF TESTS AVAILABLE TO FIT VARYING NEEDS. SURELY IT IS NECESSARY TO CONSIDER THE DIFFERENT NEEDS OF OUR SCHOOL SOCIETY AND TO USE THE PROPER INSTRUMENT AT THE PROPER TIME. IT IS INSTRUCTIVE TO CONSIDER THE CONCEPTS OF FORMATIVE AND SUMMATIVE TESTING AT THIS POINT. WHILE BY NO MEANS EXCLUSIVE OF ONE ANOTHER, WE TEND TO THINK OF CRITERION REFERENCED TESTS AS PARTICULARLY APROPOS TO THE DYNAMIC DEMANDS OF THE INSTRUCTIONAL SETTING (FORMATIVE), AND THE NORM REFERENCED TO THE LONGER SPANNED EDUCATIONAL EFFORT IN WHICH, INDEED, ONES DOES WISH TO "SUM UP" AND TO COMPARE (SUMMATIVE). IT IS A FASCINATING EXERCISE TO TRY TO FATHOM WHY CERTAIN MEASUREMENT "NEEDS" ARISE. SOMETIMES THEY ARE MERE REPRESENTATIONS OF WHAT LATER COME TO BE CALLED FADS. SOMETIMES THEY SEEM TO BE GENERATED BY BIG BUCKETS OF FEDERAL MONEY; AND SOMETIMES THEY SEEM TO BE REAL MANIFESTATION OF SOME UNDERLYING DISCONTENT. I THINK WE SEE THE LAST NOW -- ESPECIALLY WITH TEACHERS -- THE NEED TO HAVE TEST INFORMATION THAT IS THEIRS, NOT THE ADMINISTRATOR'S OR THE SCHOOL BOARD'S. THEY WANT THE

INFORMATION TO HELP THEM DECIDE WHAT TO DO NEXT IN THE INSTRUCTIONAL SETTING. IT IS HERE THAT TEST ITEMS, WELL WRITTEN, EDITED, AND ASSIGNED TO SOME CLEARLY SENSIBLE OBJECTIVES, WILL YIELD INDICATIONS OF PUPIL BEHAVIOR WHICH ARE TIMELY AND MANAGEABLE IN TERMS OF NEXT INSTRUCTIONAL STEPS. I DON'T CARE WHETHER THESE TESTS BE CALLED CRITERION REFERENCED, DIAGNOSTIC, OR WHAT -- AND NEITHER DOES THE TEACHER. CURIOUSLY, TESTS SUCH AS THESE HAVE BEEN AVAILABLE BEFORE -- AND HAVE BEEN PUBLISHING FAILURES. THERE WERE AT LEAST THREE REASONS FOR THIS I BELIEVE: (1) THEY TOOK "TOO MUCH TIME;" (2) THEY WERE "TOO COMPLICATED TO ADMINISTER AND SCORE;" (3) THEY OVERWHELMED THE TEACHER BY DEMANDING KNOWLEDGE AND TRAINING SHE DID NOT HAVE. IF WE ARE TO AVOID FAILURE AGAIN, WE SHALL HAVE TO WATCH OURSELVES CLOSELY. WE WILL HAVE TO IMPROVE AND EXTEND OUR MEASUREMENT TRAINING EFFORTS IN UNIVERSITIES AND COLLEGES AND IF CRITERION REFERENCED AND DIAGNOSTIC TESTING ARE REALLY TO HAVE AN IMPACT, IT WILL MEAN THAT THIS TRAINING SHOULD TAKE PLACE AS THE TEACHER-TO-BE IS LEARNING HOW TO TEACH READING, OR MATH, OR BIOLOGY. SUCH EFFORTS CANNOT BE LEFT TO THE TYPICAL TESTS AND MEASUREMENTS COURSE AS IT HAS BEEN TRADITIONALLY TAUGHT, FOR WE ARE TALKING HERE OF FORMATIVE TESTING - SOMETHING INEXTRICABLY BONDED TO INSTRUCTION. TO PULL THIS OFF, THE PUBLISHER WILL ALSO

HAVE TO DO HIS PART THROUGH FURNISHING FAR BETTER INTERPRETATION AND TRAINING MATERIALS THAN NOW ACCOMPANY MOST TESTS, AND THROUGH MORE INTENSIVE IN-SERVICE TRAINING OF SCHOOL PERSONNEL IN WORKSHOPS AND OTHER GROUP ACTIVITIES AT THE SCHOOL SITE. I CAN'T RESIST A FINAL THOUGHT HERE BECAUSE IT SEEMS TO ME TO BE SO ELEMENTARY AND NECESSARY. WE MUST STRIVE FOR MORE COOPERATION AND UNDERSTANDING BETWEEN PUBLISHER AND UNIVERSITY TRAINING EFFORTS. BOTH ARE SUBSTANTIAL AND NECESSARY. IT IS RIDICULOUS AND WASTEFUL THAT THEY ARE NOT MORE HELPFUL TO, AND UNDERSTANDING OF, ONE ANOTHER.

I HAVE CHOSEN NOT TO TALK ABOUT THE PROBLEMS OF CLASSIFYING OBJECTIVES; THE RELIABLE MATCHING OF ITEMS TO OBJECTIVES; AND THE INCOMPLETE STATE OF THE ART IN TERMS OF RELIABILITY AND VALIDITY ESTIMATES. OTHERS WILL TALK ABOUT THESE MATTERS, I PRESUME. OF EQUAL IMPORTANCE FOR THE SURVIVAL OF THE CRITERION REFERENCED MOVEMENT ARE STEPS NEEDED TO SEE TO IT THAT IT DOESN'T COLLAPSE OF ITS OWN WEIGHT. I SPEAK HERE OF FORCED TEACHER LABOR IN BEHALF OF INSTITUTING A ROSTER OF OBJECTIVES UNWANTED BY THAT TEACHER, OF TEACHER RELUCTANCE IN WRITING OBJECTIVES STATED TO THE POINT OF NIGGLING SPECIFICITY AND OF THE INEVITABLE REALIZATION OF THE TIME THAT IT ALL TAKES COMPARED TO REAL, AVAILABLE CLASSROOM TIME. I AM NOT AS MUCH CONCERNED FOR CRITERION REFERENCED TESTS USED FOR PROGRAM EVALUATION -- HERE THE TIME ELEMENT, BOTH IN INSTRUMENT

PREPARATION AND TESTING TIME NEEDED, PROBABLY WILL NOT BE SEEN AS TOO DEMANDING.

I SUGGEST THAT WE CONTINUE OUR CRITERION REFERENCED ENDEAVORS ON A NOTE OF OPTIMISTIC CAUTION. I WANT TO BE OPTIMISTIC BECAUSE I CAN'T BELIEVE THAT ANY PROCESS WHICH FORCES ONE TO BE CLEAR ABOUT WHERE ONE IS HEADED BEFORE ONE STARTS, CANNOT HELP BUT BE A GOOD. I WANT TO BE CAUTIOUS BECAUSE IT CAN ALSO LEAD TO OVERSPECIFICITY AND RIGIDITY. TO PARAPHRASE TRUMAN KELLEY IN A STANFORD ACHIEVEMENT TEST AUTHORIAL MEETING CIRCA 1930, "FRIENDS, JUST REMEMBER. YOU DON'T MEASURE WHAT THE SCHOOL TEACHES, YOU MEASURE WHAT THE CHILD KNOWS."